

Table 1. Data sources of variables, methods of assessment, missing data rules and details of subgroups.

Variable	Methods of assessment	Units	For multivariable analysis, missing values classified as:	Subgroups chosen for analysis
Age [years]	Self-report	Years	NA	18-29; 30-49; 50+
Sex	Self-report	NA	NA	Male; Female
Born in the UK	Self-report	NA	NA	Yes; No
Total time spent homeless	Self-report	Years	NA	<1 year; 1 year; 2-3 years; >3 years
Ever spent time in prison	Self-report	NA	No	No; Yes; Missing
Illicit drug usage	Self-report	NA	NA	None; Ever smoked heroin / crack; Ever injected drugs Highest risk group for BBVs chosen i.e. if smokes and injects then classified as injects.
Case currently smokes cigarettes	Self-report	NA	NA	Yes; No
Participant or health worker ever been concerned about drinking	Self-report	NA	No	Yes; No